Title Slide

- Personal Expense Tracker – Data Science Capstone Project
- Analyzing and Predicting Personal Spending Patterns
- Aanchal Gor
- Date: [Insert Date]

Executive Summary

- Aim: Predict potential savings based on income, expenses, lifestyle.
- Dataset: Personal Expense Tracker with 19 columns.
- Methodology: Data Collection → Data Cleaning → EDA → SQL Analysis → Predictive Modeling → Dashboard & Map.
- Key Findings: Age, Income, City Tier affect savings; Random Forest predicts well.

Introduction

- Problem: Individuals struggle to manage finances.
- Objective: Identify patterns & predict potential savings.
- Relevance: Helps budget planning.

Data Collection & Wrangling

- Data source: CSV dataset.
- Cleaning: Handle missing values, drop duplicates, convert categorical.
- Placeholder: Table screenshot of dataset head.

EDA & Interactive Visual Analytics

- Techniques: Histograms, Boxplots, Correlation Heatmaps, Pairplots.
- Insights: Income higher in City Tier 1; Age correlates with savings; Rent & Loans affect Disposable Income.
- Placeholder: Include heatmap, histogram, boxplot images.

Predictive Analysis Methodology

- Target: Potential_Savings > 5000 (binary)
- Features: Income, Age, Rent, Dependents
- Model: Random Forest
- Train/Test: 80% / 20%
- Performance metric: Accuracy & Confusion Matrix

EDA with Visualization Results

- Charts: Bar, Scatter, Pie, Pairplots
- Insights: Income distribution, spending patterns, correlations
- Placeholder: Chart screenshots

EDA with SQL Results

- Queries: Avg Income by City Tier, Total Rent by Occupation, Sum of Expenses
- Table screenshots
- Insights: Metro cities higher income, engineers/managers higher rent, food & rent dominate expenses

Interactive Map with Folium

- Map markers by city, Income & City Tier info
- Placeholder: Map screenshot or embed

Plotly Dashboard Results

- Interactive bar charts, scatter plots, pie charts
- Filter by City Tier, Occupation, Age
- Placeholder: Dashboard screenshot

Predictive Analysis Results

- Accuracy: [Insert Accuracy]
- Confusion Matrix screenshot
- Insights: Higher income & fewer dependents → higher predicted savings
- Placeholder: Feature importance plot or heatmap

Conclusion

- Key Findings: Income, Age, Rent, City Tier influence savings; dashboards/maps help understanding
- Future Work: Deploy web app, include more lifestyle variables

Creativity & Innovative Insights

- Extra Features: Interactive charts, Folium map, highlight patterns (e.g., students save less, doctors save more)
- Placeholder: Infographic of insights/trends